

METHOD AND APPARATUS FOR MANAGING DIRECT I/O TO STORAGE SYSTEMS IN VIRTUALIZATION

BACKGROUND OF THE INVENTION

5 The present invention relates generally to method and apparatus of virtualizing network storage. More particularly the present invention relates to a method and apparatus for virtualizing network storage and permitting access to the network storage separate from virtualization apparatus.

 A conventional information processing system stores information in a
10 storage directly connected to a computer system. An access to information stored in the storage is permitted only through the use of the directly connected computer. Accordingly, the other computer systems need to access the data via the directly connected computer. In recent years, the development of network technologies and a dramatic increase in the amount
15 of information to be stored promote separation of a computer system for processing information and a storage system for storing information. Thus, storage can be connected to the network and has become capable of being shared by a plurality of computer systems. Such networked storage is referred to as a network storage.

20 Examples of such network storage include a storage area network (SAN) storage that is connected by a SAN and provides a block access, a network attached storage (NAS) that is connected for example to an IP network and provides file access, and a Web storage that offers an access according to an interface HTTP protocol or its extended protocol for Web
25 accesses.

As network storages become widely used, a system administrator needs to manage a plurality of network storages connected to the network. When the amount of data to be stored exceeds a maximum value for the existing network storage, for example, it is necessary to add a new network storage and reconfigure the system by re-distributing data or the like. A drastic increase in the need for storage capacities forces repeated system reconfigurations, thus increasing system management costs.

In order to conserve system management costs, a technology for virtualizing storages, specifically consolidating the storage resources as "one big capacity pool" is mandatory so that a plurality of network storages can be viewed as a single storage from the computer system and the entire system is not affected by addition of new devices. Open heterogeneous environments need such technology to manage those resources more efficiently. Various systems have been developed and proposed for virtualizing storages.

However, the conventional configuration of a virtualization box for virtualizing storages is not particularly good being that a customer can not access their data via the host channels for connecting the host to the storage device for other purposes such as reporting, verifying or backing up. In this configuration the customer is required to use one of the ports of the virtualization box.

SUMMARY OF THE INVENTION

The present invention provides a method and apparatus for virtualizing network storage and permitting access to the network storage separate from virtualization apparatus.

Particularly the present invention provides a network storage virtualization method and apparatus in a network storage system having a plurality of network storage devices. According to the present invention a client connected via a network is allowed to access the network storage devices as one virtual network storage system. In addition the client is permitted to access the network storage devices separate from the virtualized network storage system for other purposes such as reporting, verifying or backing up.

10 BRIEF DESCRIPTION OF THE DRAWINGS

The foregoing and a better understanding of the present invention will become apparent from the following detailed description of example embodiments and the claims when read in connection with the accompanying drawings, all forming a part of the disclosure of this invention. While the foregoing and following written and illustrated disclosure focuses on disclosing example embodiments of the invention, it should be clearly understood that the same is by way of illustration and example only and the invention is not limited thereto, wherein in the following brief description of the drawings:

Figs. 1A and B are diagrams for explaining the basic system configuration of the present invention with replication;

Fig. 2 shows an example of the internal configuration of a storage system having a storage controller;

Fig. 3 illustrates another example of the internal configuration of a storage system having an intelligent switch;

Fig. 4 is a flowchart of the process performed to configure an access permission of a volume from outside of a storage system controller;

Fig. 5 illustrates the steps of a process to access a volume for purposes such as reporting in the read only mode; and

5 Fig. 6 illustrates the steps of a process to access a volume for purposes such as reporting in the read/write mode.

DETAILED DESCRIPTION OF THE INVENTION

Examples of the consolidation of storage resources are shown in Figs.
10 2 and 3.

As per Fig. 2, the storage controller 201 includes a front end controller 2010 for interfacing with a channel of a host, back end controller 2011 for interfacing with a storage device 202 via disk interface (I/F) 108-1 and 108-2, an external storage controller 2012 for interfacing with a disk array 203,
15 shared memory 2013 for caching data including control data transferred between the host and storage device 202 and the disk array 203 and an interconnection apparatus 2014 for interconnecting the front end controller 2010, back end and external storage controllers 2011 and 2012, respectively and shared memory 2013. Back end controller 2011 interfaces to the primary
20 volume 103 and external storage controller 2012 interfaces to local secondary volume 109. It should be noted that the primary volume 103 and the local secondary volume 104 are logical volumes. The primary volume 103 is defined on a plurality of disk devices 2020 included in the storage device 202.

The local secondary volume 104 is defined on a plurality of disk drives
25 2030 included in the disk array 203 which includes a host interface (I/F) 2031

for interfacing with the storage controller 201, a disk I/F 2032 for interfacing with the disk drives 2030, a shared memory 2033 for caching data including control data transferred between the storage controller 201 and disk drives 2030, and an interconnection apparatus 2034 for interconnecting the host I/F 2030, disk I/F 2032, and shared memory 2033.

As illustrated in Fig. 3 an intelligent switch (SW) 301 in place of the storage controller 201 as per Fig. 2 is provided having a front end (host) controller 3010 for interfacing with a channel of a host, a back end controller 3011 for interfacing with the storage device 202 and the disk array 203 and an interconnection apparatus 3014 for interconnecting the front end controller 3010 and back end controller 3011. The back end controller 3011 interfaces to the primary volume 103 and the local secondary volume 109. The front end controller 2010 includes mapping tables 30101 having stored therein mapping information for mapping virtual volumes to the storage device 202 or the disk array 203. Particularly the mapping tables 30101 set relationships between virtual volume block address and actual block address on the storage device 202 or the disk array 203. As described above the primary volume 103 is defined on a plurality of disk devices 2020 included in the storage device 202 and the local secondary volume 104 is defined on a plurality of disk drives 2030 included in the disk array 203.

The various storage resources such the storage device 202, disk array 203, or any other such storage such as just a bunch of disks (JBOD), etc., are connected to a virtualization box such as the storage controller 201 or the intelligent SW 301, which can organize the various storages. In this case the ports of the storage device 202 and the disk array 203 are connected to the

storage controller 201 or the intelligent SW 301 and these ports are all under the control of the storage controller 201 or the intelligent SW 301. Storage device 202, the disk array 203 or any other such storage device have their own high-speed host channels such as Fibre channel or Giga Ethernet. If all
5 of the ports of the storage device 202 and the disk array 203 are connect to the storage controller 201 or the intelligent SW 301, then there are no connections to the host channels of the storage devices available for other uses.

Fig. 1A illustrates the basic system configuration of the present
10 invention with replication. As per Fig. 1A, a first storage system 101 is provided and includes a primary volume (P) 103, a local secondary volume (S) 104 which is a local mirror or snap shot of the primary volume at a certain point in time, and other volumes 105. Primary volume 103 is provided as part of the storage system 101 via a storage device 106 and the local secondary
15 volume 104 and the other volumes 105 are provided by a storage device 102. The storage system 101 also includes a controller 107 for controlling the storage system 101 that interfaces with each of the storage devices 106 and 102 via a disk interface (I/F) 108-1 and interface (I/F) 108-2 respectively. The storage system 101 is connected to a host 109 so as to receive input/output
20 (I/O) commands, for example, write I/O commands from the host 109.

The storage device 102 also includes a host I/F 111 which allows the storage device 102 to interface, for example, with a host 120 via a storage area network (SAN) 110.

I/F 108-2 is similar to the host I/F 111. I/F 108-2 works as a target port
25 for the storage device 102 and controller 107 recognizes I/F 108-2 as a target

port for access to the local secondary volumes 104 and possibly the other volumes 105. Thus controller 107 recognizes I/F 108-2 in the same manner as host 120 recognizes host I/F 111 as a target port for access to volume 104.

Thus, according to the present invention, the storage controller 107 provides virtualization functions with respect to the storage devices 106 and 102 thereby organizing the storage devices 106 and 102 to appear as a single storage when accessed by the host 109. However, the present invention provides apparatus so as to permit access to the storage such as, for example, provided by storage device 106 separate from the virtualization function being performed by the controller 107. The present invention accomplishes such by providing the storage device 102 with the host I/F 111 so as to permit the host 120 via the SAN 110 to access the local secondary volume 104 for reporting verifying backing up or any other such functions separate from the virtualization functions provided by the controller 107. Access to the local secondary volumes 104 are provided by the storage device 102 in a manner such that access to the other volumes 105 are not permitted. Particularly, access to the other volumes 105 in the storage device 102 are locked out in a manner so as to only permit access from the storage system 101 via the I/F 108-2.

Fig. 1B illustrates a configuration similar to that of Fig. 1A with the exception that the local secondary volume 104 and the other volumes 105 are provided by the storage device 102 which is connected to the controller 107 via a wide storage area network (WSAN) 121 and a host I/F 111-1. Thus, the principle is still the same as that described above wherein access to the local secondary volume 104 is provided to the host 120 via the SAN 110 and the

host I/F 111-2 separate from the virtualization functions provided by the controller 107. As described above, access to the other volume 105 by the host 120 for any device other than the storage system 101 is not permitted and is specifically locked out.

5 According to the present invention, access to the other volumes 105 of the storage device 102 is locked out using at least two different methods. A first method is not expose the other volumes 105 to the host I/F 111 or the host I/F 111-2 which interface with the host 120 via the SAN 110. Another method is to filter I/O accesses to the storage device 102 using logical unit
10 number (LUN) masking or a LUN security function and the storage device 102. This method allows for a registered node having a pre-designated world wide name (WWN) to access the other volumes 105. Still yet another method is to mask the LUN in the host bus adapter (HBA) level. The HBA is an I/O adapter that sits between the computer's bus and the fiber channel loop and
15 manages the transfers of information between the two channels. The HBA can map specific LUNs for filtering. Thus, the key point according to the present invention as is clear from the above is to make the function of permitting access to the secondary volume separate from the virtualization in alignment with the user's definition, pair status or policy of direct I/O to the
20 secondary volume from outside of the storage system control.

 Another manner of permitting access to the secondary volume 104 of the second storage device 102 is to set the second storage device 102 to permit direct I/O access from, for example, the host 120 via the SAN 110. In order to accomplish such the replication function and set up access control
25 must be configured. The flowchart illustrated in Fig. 4 shows the process of

the steps to configure the access permission of the secondary volume 104 from outside of the controller 107.

As illustrated in Fig. 4 the access control for the secondary volume 104 must be set to, for example, read only mode before creating the local or remote mirroring pair (step 401). Thereafter, the mirroring pair between primary volume 103 and secondary volume 104 is created (step 402). So as to establish a mirror or snap shot of the primary volume in the secondary volume 104 at a certain point in time a complete initial copy of the primary volume 103 is taken and stored in the secondary volume 104 (step 403).

10 Thereafter, the second storage device 102 receives a suspend request (step 404) and places the mirroring pair into the suspend status so as to permit the setting of the access control according to the user choice (step 405). If the user chooses the read only mode for the secondary volume 104 in the second storage device 102 such as for reporting purposes, the secondary volume 104

15 is set to the read only mode for I/Os from outside of the storage controller 107 (step 406). Thereafter, the mirroring pair receives a re-synchronization request so as to re-synchronize the mirroring pair when in the read only mode (step 407). However, if the user chooses the read/write mode for the secondary volume 104 then the read/write mode in the secondary volume 104

20 of the second storage device 102 is set (step 408). Thereafter, the mirroring pair receives the re-synchronization request so as to re-synchronize the secondary volume 104 with the primary volume 103 using the initial copy (steps 409 and 410). However, if the user can make sure that no write request has been performed the user can specify the specific option for re-

25 synchronization thereby avoiding use of the initial copy (step 411).

Fig. 5 illustrates the steps of the process to be performed to permit access of the secondary volume for reporting purposes in a read only mode. As per the flowchart in Fig. 5 synchronization of the mirroring pair, namely, the primary volume 103 and the secondary volume 104 which are set in the synchronous mode is suspended (step 501). Thereafter, access to data on the secondary volume 104 is permitted via the host I/F 111. Once the access is completed synchronization between the remote mirroring pair is reestablished (step 503).

Fig. 6 illustrates the details of the steps to be performed to permit access of the secondary volume 104 when it is set in the read/write mode. As per the flowchart of Fig. 6 synchronization of the mirroring pair, namely the primary volume 103 and the secondary volume 104 being set in the read/write mode, is suspended (step 601). Once suspended access to the secondary volume 104 is permitted via the host I/F 111 (step 602). Thereafter, re-synchronization is performed in the mirroring pair and if no specific option is selected (step 603) an initial copy is made of the primary volume 103 (step 604). However, alternatively once access to the secondary volume 104 is completed re-synchronization of the mirroring pair is conducted with a specific option such as, for example, no INIT (step 605). Thereafter, the usual re-synchronization process is started (step 606).

Therefore, according to the above, the present invention provides a network storage visualization method and apparatus in a network storage system having a plurality of network storage devices wherein a client is allowed to access the network storage devices as one virtual network storage system and is permitted to access the network storage device separate from

the virtualized network storage system. According to the present invention such is accomplished by providing a storage device such as that illustrated in Figs. 1A and B of the present application, such storage device having a host I/F 111 to which a host 120 via, for example, a SAN 110 can access the
5 secondary volume 104 of a mirroring pair without use of the virtualization functions provided by the storage controller 107.

While the invention has been described in terms of its preferred embodiments, it should be understood that numerous modifications may be made thereto without departing from the spirit and scope of the present
10 invention. It is intended that all such modifications fall within the scope of the appended claims.